Noise reduction by Spectral Subtraction for Speech Recognition

[1] Shital S. Joshi, [2] Vaishali D. Bhagile

[1]Research Scholar, [2]Assistant Professor,
[1]Department of Computer Science & IT,
[1] Dr. Babasaheb Ambedkar Marathwada University, Aurangabad 431002, Maharashtra, India,
2Deogiri Institute of Technology and Msnsgement Studies, Aurangabad-431005, Maharashtra, India.

Abstract : This paper is an attempt to explore noise reduction by spectral subtraction method. We have worked on two databases. Database 1 is recorded in a noisy environment and database 2 is recorded in noise free environment. Spectral subtraction method is used to reduce the noise and enhance speech signal. Experiment gives good results on both of the datasets. Output speech signal is more clean and clean than noisy speech signal. Also we got SNR >20 dB.

IndexTerms **- Spectral Subtraction, Noise Reduction, Speech Enhancement etc.**

I. INTRODUCTION

A. Speech

Speech is a very natural form of communication, massage passing and express our ideas, thoughts and emotions within human beings. Speech recognition is very interesting area of research. It deals with the process of converting speech signal to text. Essentially speech signal must be pure to achieve good recognition rate and higher accuracy. Speech samples that are recorded by microphone is generally infected by noise originating from various sources [1].

B. Noise

Generally, Noise is a general term which is used for unwanted element present in the desired signal [2]. Noise can be defined as an unwanted signal that interferes with the communication or measurement of another signal [3]. Noise is assumed to be long term stationary and its spectrum estimated from frames that do not contains speech silent frames contain background noise [4]. Noise can be classified as adaptive, background, random, additive, natural, musical, multi-talker, babble, non-stationary, ambient, white Gaussian etc.

C. Effect of noise on speech

Noise influences the accuracy of Automatic Speech Recognition system. It has very crucial impact on the performance of speech signal and its recognition systems. The performance of automatic speech recognizer (ASRs) degrade rapidly in the presence of noise and other distortions [5]. For large scale & real-world applications noise processing is becoming highly important as speech recognition needs to work in highly demanding acoustic environment rather than past [6].

D. Speech Enhancement

Noisy speech is processed to minimize the effect of noise in speech signals[7]. Speech enhancement is referred as to improve quality or intelligibility of speech signal. Speech enhancement methods like-

1. Spectral subtraction
2. Modified Spectral Subtraction
3. Least Mean Square
4. Linear and Butterworth filters

All these are single channel speech enhancement methods. Speech enhancement reduces background noise, improve SNR and assumes stationary noise that noise is more stationary than speech [8].

E. Features of Spectral Subtraction Method

- This technique is based on direct estimation of short term spectral magnitude [9].
- Spectral Subtraction speech enhancement technique perform better due to good speech reconstruction quality [9].
- The Noise corrupted speech is processed by Spectral Subtraction method to get processed enhanced speech. Spectral Subtraction is a popular frequency domain method to reduce the effect of additive uncorrelated noise in a single [10].
- The key advantage of this method of speech enhancement is that it is easy to implement.
- In this method, noise magnitude is estimated by the input waveform itself. And therefor no need to do spectral analysis and design bandpass filters for each waveform.

Many audio editing tools (e.g. PRATT, audacity etc.) also provide facility to reduce noise from audio by using Spectral Subtraction method.

II. SPECTRAL SUBTRACTION METHOD

Spectral Subtraction method was proposed by Boll is a popular speech enhancement technique through noise reduction. In Spectral Subtraction, an estimate of speech signal is simply obtained by subtracting a pre-estimated noise spectrum from observed one. It is single channel enhancement method for enhancing speech degraded by additive background noise [11].Spectra Subtraction needs to be done very carefully to avoid any speech distortion. If too much is subtracted, then some speech information might be removed, whereas too little is subtracted, then much of interfering noise remains [12].

III. BASIC PRINCIPLE OF SPECTRAL SUBTRACTION METHOD

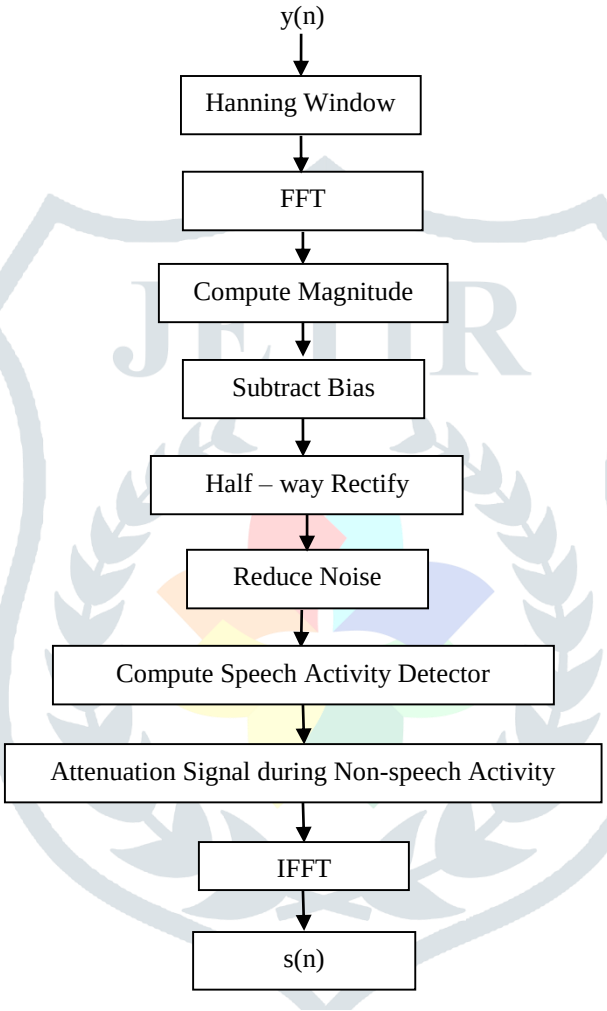

Fig. 1. : Spectral Subtraction Method [13]

Assume that y(n), is the noise-corrupted input signal. It is composed of the clean signal x(n) and the additive noise signal, d(n).

$$y(n) = x(n) + d(n) \qquad 0 \le n \le N-1 \quad [14]$$

This can be represented in frequency domain as follows:

$$Y(k) = X(k) + D(k) \qquad 0 \le k \le N-1 \quad [14]$$

Taking Fourier transform, the power spectrum expectation is :

$$E[|Y(k)|^2] = E[|X(k)|^2] + E[|D(k)|^2] \qquad 0 \le k \le N-1$$

Average signal of frames without speech that means noise is estimated as $E[|D(k)|^2]$ which is denoted by λ_k^2.And this noise is subtracted from speech signal.

$$|S(n)| = \begin{cases} |Y(k)|^2 - \lambda_k^2 & |Y(k)|^2 \ge \lambda_k^2 \\ 0 & else \end{cases} \quad [14]$$

Finally, Inverse Fourier transformation is done to get clean speech.

IV. EXPERIMENTAL SETUP

We have performed noise this experiment on two datasets. Dataset1 is Marathi speech numeric data which is recorded in noisy environment. Marathi numbers were recorded from 100 speakers. Each digit was repeated 5 times. Isolated Marathi Numbers were recorded in a mono channel using PRATT software and stored in .wav file. Total vocabulary size of database 1 is 5000[15].

Dataset 2 is Marathi Numeric, alphabetic data recorded in a closed room. Sampling frequency 16000 Hz and bit rate was 16 bits. Connected words were recorded in mono channel. Total Database2 size is 300 wav files.

V. RESULTS AND DISCUSSION

The quality of any noise reduction method can be measured by viewing its refinement in SNR i.e. Signal-To-Noise Ratio. The performance of algorithm can also be evaluated by object measures like Mean Square Error (MSE), Root Mean Square Error (RMSE) & Normalized Root Mean Square Error (NRMS). Among all these SNR is most widely and popular method to measure the quality of speech. It is ratio of signal to noise power in decibels(dB).

$$\mathrm{SNR} = \log_{10}\left(\frac{\sigma_x^2}{\sigma_d^2}\right)$$

Where σ_x^2 is the mean square of speech signal

σ_d^2 is mean square difference between original and reconstructed speech[16].

SNR estimation can improve algorithms for speech enhancement, speech detection and speech recognition. Knowledge of SNR makes it easier to compensate for the effects of noise [17].

We have measured signal-to-noise (SNR) ratio by using in-built function snr(). This function is provided by Digital Signal Processing toolbox in Matlab. Normally SNR>20 is considered as a good signal in speech processing.

The results are explored in following figures. Figure 2 illustrates the result of noise removal method in signal as well as in a spectrogram. We can see that the noise frequencies are present in spectrum of noisy signal whereas noisy frequencies are absent in output speech signal. Frequency Spectrum of output signal is lighten than that of input signal. Here we get SNR as 22.5519 dB.

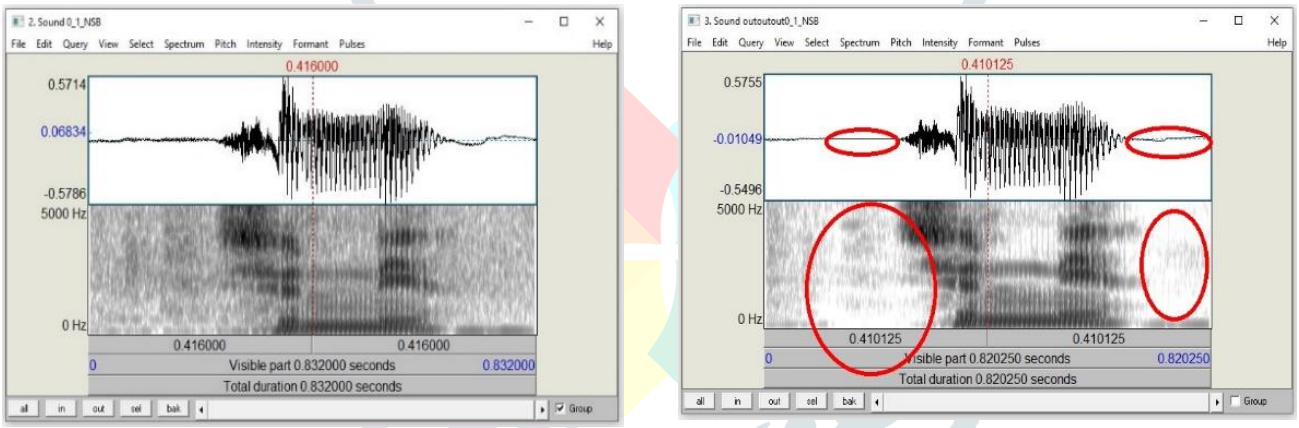

Fig. 2. Noisy Speech Signal and Noise Removed Output Speech signal of Dataset 1

Figure 3 explore results of the experiment performed on dataset 2. Speech signals with corresponding frequency spectrum are showed in figure 3. Noisy speech have dark noise frequencies present in its spectrogram whereas output signal have only dark speech signals that means noise frequencies are removed from it.

SNR ratio obtained for this sample is 21.0965 dB

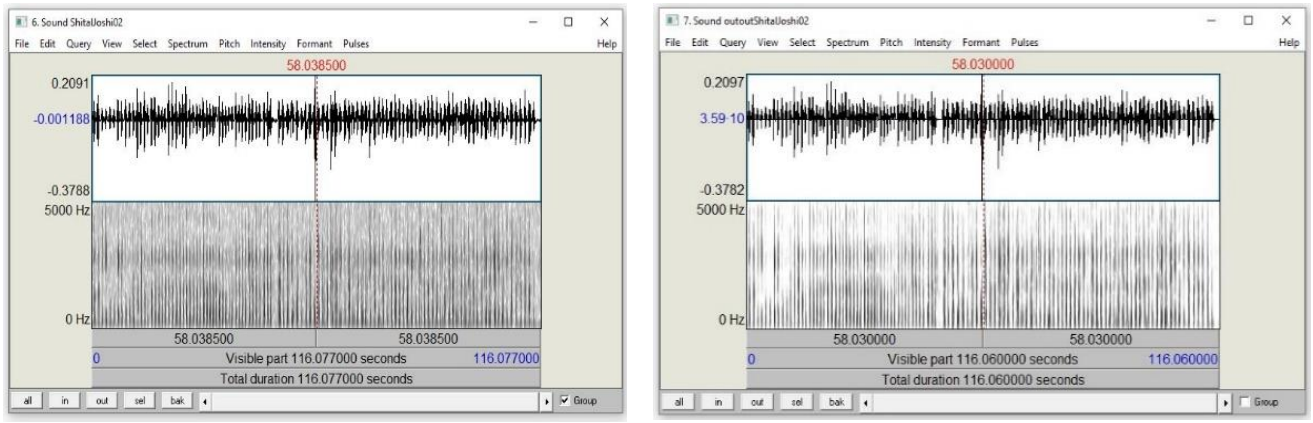

Fig. 3. Noisy Speech signal and noise Reduced Output Speech Signal of Dataset 2

VI. CONCLUSION

Noise is efficiently removed from noisy speech signal for isolated and connected Marathi speech databases. Enhanced speech wave files are more audible and clear than input noisy speech wave files. Enhanced speech can produce better results than noisy speech signals.

V. ACKNOWLEDGMENT

Authors would thanks to P. P. Shrishrimal et. al. for sharing their dataset of isolated Marathi numerals. Authors would also like to thank the speakers for sparing their time for recording of Dataset 2. Authors would also like to thank University authorities for providing infrastructure to carry out the research.

REFERENCES

[1] **Shrawankar, U., & Thakare, V.: Noise estimation and noise removal techniques for speech recognition in adverse environment. In *International Conference on Intelligent Information Processing* (pp. 336-342). Springer, Berlin, Heidelberg. (2010).**

[2] **Garg, K., & Jain, G.: A comparative study of noise reduction techniques for automatic speech recognition systems. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 2098-2103). IEEE. (2016).**

[3] **Sathya, J., Victor, S. P.: Noise Reduction Techniques and Algorithms For Speech Signal Processing. *International Journal for Research in Science Engineering & Technology*, *2*(11), 34-41. (2015).**

[4] **Srinivasan, S., Wang, D.: Robust speech recognition by integrating speech separation and hypothesis testing. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.* (Vol. 1, pp. I-89). IEEE. (2005).**

[5] **Gong, Y.: Speech recognition in noisy environments: A survey. *Speech communication*, *16*(3), 261-291.(1995).**

[6] **Li, J., Deng, L., Gong, Y., Haeb-Umbach, R.: An overview of noise-robust automatic speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *22*(4), 745-777. (2014).**

[7] **Garg, K., & Jain, G.: A comparative study of noise reduction techniques for automatic speech recognition systems. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 2098-2103). IEEE. (2016).**

[8] **Singh N., Gupta D.: A Review Paper: Speech Enhancement Using Different Spectral Subtraction Algorithms. IJLTMAS, Vol III, Issue IV,ISSN 2278-2540. (2014).**

[9] **Pawar A.P., Chaiudhari K.B.: Enhancement of Speech in Noisy Conditions, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 2, Issue 7. (2013).**

[10] **Venkateswarlu, S. C., Reddy, A. S. R., Prasad, K. S.: Speech Enhancement using Boll's Spectral Subtraction Method based on Gaussian Window. *Global Journal of Researches in Engineering: Electrical and Electronics Engineering*, *14*(6). (2004).**

[11] **Pawar, A. P., Choudhari, K. B., & Joshi, M. A. (2012). Review of single channel speech enhancement methods in spectral domain. *Int. J. Appl. Eng. Research*, *7*, 1961-1966.**

[12] **Loizou, P. C.: *Speech enhancement: theory and practice*. CRC press.2nd Edition. (2013).**

[13] **Boll, S. (1979). Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on acoustics, speech, and signal processing*, *27*(2), 113-120.**

[14] **Gao, L., Guo, Y., Li, S., & Chen, F.: Speech enhancement algorithm based on improved spectral subtraction. In *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems* (Vol. 3, pp. 140-143). IEEE. (2009).**

[15] **Shrishirmal, P. P., Deshmukh, R. R., Waghmare, V. B., Borade, S., Janse, P. V., & Janvale, G. B. Development of Marathi Language Speech Database from Marathwada Region.**

[16] **Nemade, M. U., & Shah, S. K. (2013). Performance comparison of single channel Speech enhancement techniques for personal Communication. *for International Journal of Innovative Research in Computer and Communication Engineering*, *1*(1).**

[17] **Kim, C., & Stern, R. M. (2008). Robust signal-to-noise ratio estimation based on waveform amplitude distribution analysis. In *Ninth Annual Conference of the International Speech Communication Association*.**